15.1 A Fully Integrated Multi-CPU, GPU and Memory Controller 32nm Processor

Marcelo Yuffe, Ernest Knoll, Moty Mehalel, Joseph Shor, Tsvika Kurts

Intel, Haifa, Israel

This paper describes the 32nm Sandy Bridge processor that integrates up to 4 high performance Intel Architecture (IA) cores, a power/performance optimized graphic processing unit (GPU) and memory and PCIe controllers in the same die. The Sandy Bridge architecture block diagram is shown in Fig. 15.1.1 and the floorplan of a four IA-core version is shown in Fig. 15.1.2. The Sandy Bridge IA core implements an improved branch prediction algorithm, a micro-operation (Uop) cache, a floating point Advanced Vector Extension (AVX), a second load port in the L1 cache and bigger register files in the out-of-order part of the machine; all these architecture improvements boost the IA core performance without increasing the thermal power dissipation envelope or the average power consumption (to preserve battery life in mobile systems). The CPUs and GPU share the same 8MB level-3 cache memory. The data flow is optimized by a high performance on die interconnect fabric (called "ring") that connects between the CPUs, the GPU, the L3 cache and the system agent (SA) unit that houses a 1600MT/s, dual channel DDR3 memory controller, a 20-lane PCIe gen2 controller, a two parallel pipe display engine, the power management control unit and the testability logic. An on die EPROM is used for configurability and yield optimization.

The modular ring interconnect enables the 4-core die to be easily converted into a 2-core die by "chopping" out two cores and two L3 cache modules as described in Fig. 15.1.2. Additional optimizations can be done by reducing the number of execution units of the GPU or by reducing the L3 cache size. Sandy Bridge is offered in three different C4 package types, PGA for mobile computers, LGA for desktop systems and BGA for small form factor systems. For each die flavor an optimized package stack-up was developed with 10 layers for the highest power consumption chips and 8 or 6 layers for the smaller dies.

The die is powered by 6 different power planes. The IA cores and the L3 cache share the same power plane. Power gates were uniformly spread in the cores, enabling the power control unit (PCU) to shut off any core independently. The L3 cache also incorporates power gates; this allows keeping the majority of the cache in a low power data retention mode while only the section of the cache that is accessed is fully connected to the power supply. This embedded power gate approach simplifies the package design reducing the number of package layers. The GPU is connected to its own power plane allowing independent power voltage optimization. The other power planes are used for the SA, I/O and analog circuitry. The voltage of all these power planes is controlled by the PCU using serial VID (SVID), a dedicated low frequency serial bus that connects the PCU with the external voltage regulators in a daisy chain manner.

One of the challenges of sharing the same power plane between the cores and the L3 cache is that the minimum voltage needed to keep the L3 cache data may limit the minimum operating voltage of the cores, increasing the overall average power of the system. Previous processors solved this problem by connecting the L3 cache to a separated higher voltage power plane; however this approach considerably increases the power dissipated by the L3 cache itself and taking into account Sandy Bridge implements 3MB, 4MB or 8MB of L3 cache capacity (depending on the chip configuration) the power dissipated by the cache memory accounts for a big portion of the overall power consumption of the die. Several circuit and logic design techniques have been developed to minimize the Vccmin of the L3 cache and the register files of the chip to bring it to a lower level than the core logic. Figure 15.1.3 shows the Vccmin distribution of the cache base-line and its improvement in Sandy Bridge. Fig 15.1.4 illustrates one of these techniques in a register file. Fabrication variations may cause RF write-ability degradation at low voltages (for example for wafers where T_P comes out stronger than T_N); this technique weakens the memory cell pull up device effective strength solving the low voltage write-ability issue caused by a too strong T_P device in the memory cells. The effective size of the shared PMOS is set during production testing by enabling any combination of the three parallel transistors T_1, T_2 and T_3.

The processor clocking architecture shown in Fig. 15.1.5 includes 13 PLLs driving independent clock domains for the core and cache slices, the GPU, SA as well as the four independent I/O regions [1,2]. Most of the PLLs use a 100MHz reference clock provided by the external system clock generator that is integrated into the Platform Control Hub (PCH) chip. The reference for the slices and GPU PLLs is generated by a dedicated RCLK PLL to minimize the clock skew over the different power plane domains. The SA PLL generates the clock needed by the SA logic as well as the clock for the display engine and the 133MHz reference clock for the main memory system. This clocking architecture ensures high synchronicity between the independent clock domains at the BCLK reference edges even when the domains are operating at different frequencies; to ensure data transfer between different clock domains with minimum latency low jitter PLLs (long term jitter $\sigma < 2$ps) are used. The clock distribution ensures low clock power and good clock skew performance by a combination of vertical clock spines and embedded clock compensators controlled by on die compensation state machines. Figure 15.1.6 shows the clock distribution in one of the slices. The measured clock skew is 16ps, with on-die clock compensators canceling the within die random variations.

Temperature control is extensively used in modern processors to maximize performance. When the die is hot frequency is lowered, when the die is cold the PCU takes advantage of the thermal head room to increase the frequency of the CPU. Temperature information is also used to control the system fan and to shut the CPU down in case of a catastrophic thermal event. Sandy Bridge has taken this method one step further by using two different types of thermal sensors. The first is a diode-based thermal sensor described in [3]. This sensor compares the diode voltage, which has a negative temperature coefficient to a reference voltage to output the temperature. This sensor functions over the full temperature range of operation, providing information for throttling, catastrophic function and fan regulation. Due to die area considerations there is only one such sensor per core (sensors are also located in the GPU and in the SA). Sandy Bridge has also implemented a miniaturized CMOS-based thermal sensor [4] that has a substantially reduced area (5100um^2) compared to the diode sensor, but has a more limited temperature range, being accurate between 80-100°C; since this sensor is small it can be placed at several locations inside the core providing an accurate picture of the core hot spots.

Sandy Bridge introduces the Generic Debug eXternal Connection (GDXC), a debug bus that allows monitoring the traffic between the IA cores, GPU, caches and SA on the processor internal ring. GDXC allows chip, system or software debuggers to sample ring data traffic as well as ring protocol control signals and drive it to an external logic analyzer in a packet manner through a dedicated on-package probe array. A post processing software is used to recover and analyze the data.

The Sandy Bridge thermal dissipation power (TDP) ranges from 17W to 45W for a 2-core and a 4-core mobile part and all the way to 95W for a high-end desktop part. The IA cores and GPU are powered from independent 0.7V to 1.15V variable voltage power supply sources, all controlled by the SVID bus. The DDR3 interface uses a 1.5V power plane while the PCIe interface uses a 1.05V power plane. The die photo is shown in Fig. 15.1.7.

References:

[1] Rusu, Stefan et al.: A 45nm 8-Core Enterprise Xeon Processor, ISSCC-2009

[2] Fayneh, E, Knoll, E: Clock Generation and Distribution for Intel Banias Mobile Microprocessor, VLSI Circuits Symposium, 2003

[3] Duarte, D., Geannopoulos, G., Mughal, U., Wong, K.L. and Taylor, G., "Temperature Sensor Design in a High Volume Manufacturing 65nm CMOS Digital Process", Proceedings of the 2007 IEEE Custom Integrated Circuits Conference, (CICC '07), pp. 221-224, Sept. 2007.

[4] K. Luria and J. Shor, "Miniaturized CMOS Thermal Sensor Array for Temperature Gradient Measurement in Microprocessors", Proceedings of the IEEE 2010 International Symposium of Circuit and Systems, Paris France, May 30 2010.

 978-1-61284-302-5/11/$26.00

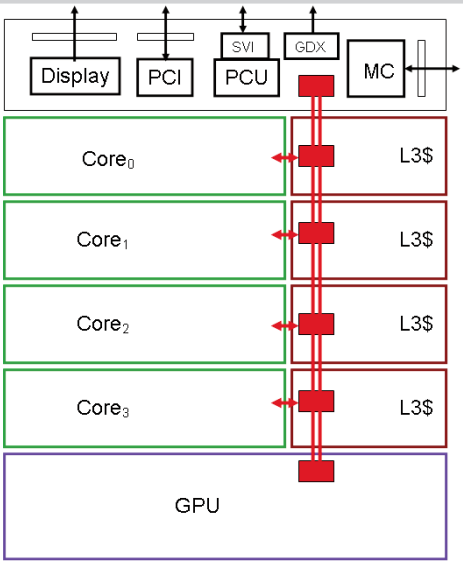

Figure 15.1.1: Sandy Bridge block diagram.

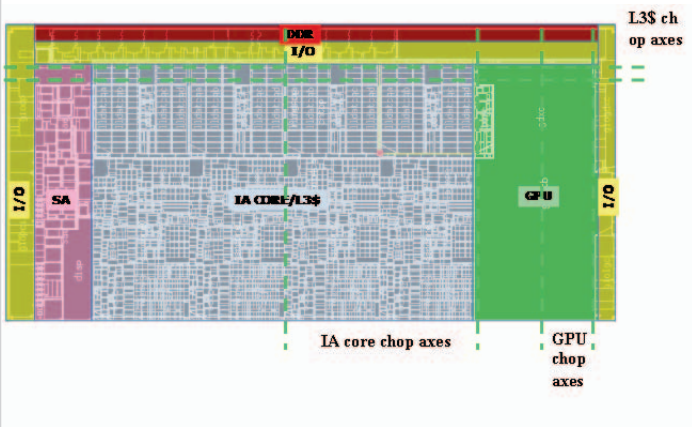

Figure 15.1.2: Sandy Bridge floorplan, power planes and choppability axes.

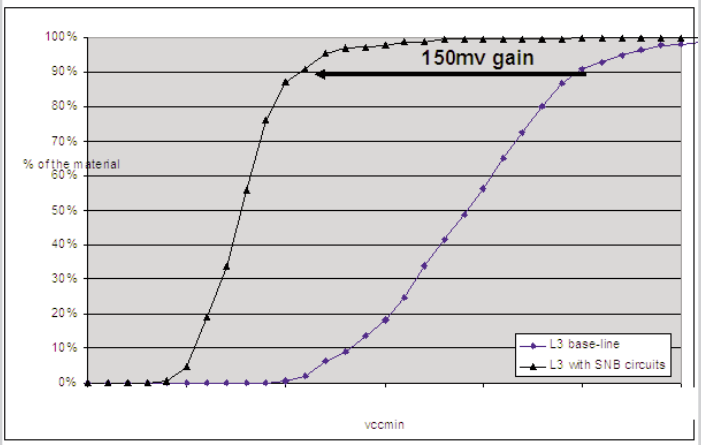

Figure 15.1.3: Vccmin improvement after applying SNB Vccmin reduction circuits.

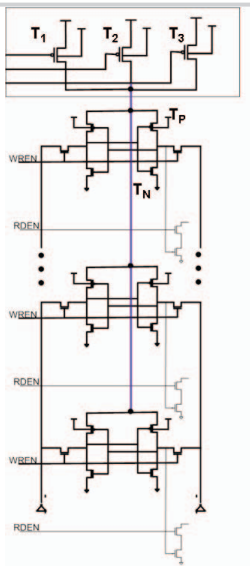

Figure 15.1.4: Register file shared strength control PMOS devices.

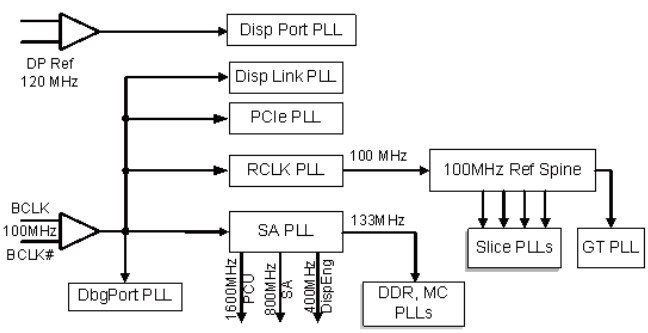

Figure 15.1.5: Sandy Bridge clock generation block diagram.

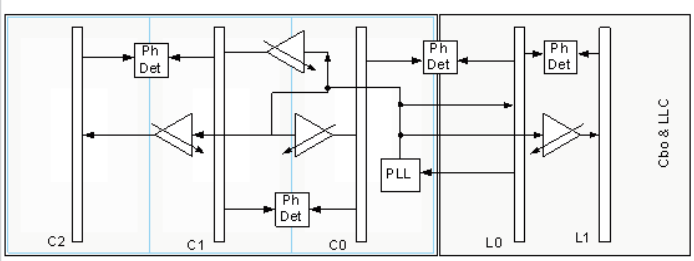

Figure 15.1.6: Slice clock distribution block diagram.

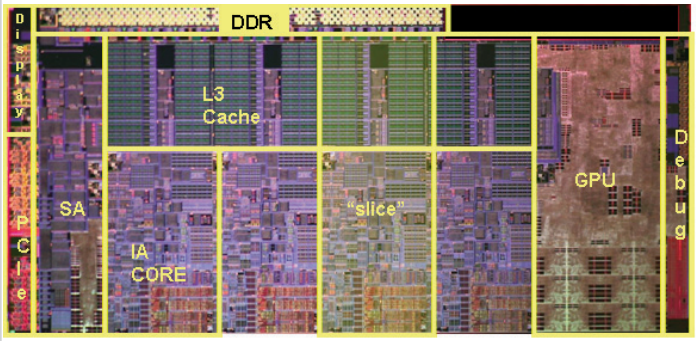

Figure 15.1.7: Sandy Bridge die photo.

978-1-61284-302-5/11/$26.00